تكنولوجيا

سمّاعات ذكية تترجم لعدة متحدثين في وقت واحد



النظام الجديد يمكن توسيعه ليشمل نحو 100 لغة (جامعة ماكاو)

القاهرة: «الشرق الأوسط»

نُشر: 14:33-11 مايو 2025 م . 14 ذو القِعدة 1446 هـ

طوّر باحثون من جامعة واشنطن الأميركية نظام سماعات رأس ذكية يمكنها ترجمة حديث عدة أشخاص في الوقت نفسه، مع الحفاظ على نبرة صوت كل متحدث واتجاه صوته بدقة ثلاثية الأبعاد، حتى أثناء حركته.

وأوضح الباحثون أن النظام، الذي أُطلق عليه «الترجمة الصوتية المكانية» قد يُحدث نقلة نوعية في مجال الترجمة الفورية. ونُشرت النتائج، الجمعة، في دورية «جمعية الحوسبة الأميركية (ACM)».

وعلى الرغم من ظهور تقنيات واعدة في مجال الترجمة الفورية مؤخراً، فإن أياً منها لم يوفر حلاً فعّالاً في الأماكن العامة. فعلى سبيل المثال، تعمل نظارات شركة ميتا الجديدة فقط عند وجود متحدث واحد، وتعرض الترجمة الصوتية الآلية بعد انتهاء المتحدث من الكلام. لكن الفريق المصمم للنظام الجديد ابتكر حلاً يُترجم كلام عدة متحدثين في الوقت نسه مع الحفاظ على الاتجاه وخصائص كل صوت.

ويعتمد النظام على سماعات مانعة للضوضاء مُزوّدة بميكروفونات، وتعمل خوارزميات النظام المستندة إلى الذكاء الاصطناعي على فصل أصوات المتحدثين وتتبعهم داخل المكان، ثم ترجمة كلامهم وإعادة تشغيله بصوت يشبه صوتهم الأصلي مع تأخير بسيط يتراوح بين 2 و4 ثوانٍ.

ويتميز النظام بثلاثة ابتكارات رئيسية، أولها قدرته على اكتشاف عدد المتحدثين تلقائياً داخل أي مساحة مغلقة أو مفتوحة، من خلال مسح محيطي بزاوية 360 درجة، وترجمة الكلام مع الحفاظ على تعبيرات الصوت ومستوى ارتفاعه، باستخدام معالجات داخلية مثل شريحة (Apple M2)، دون الاعتماد على الحوسبة السحابية لحماية الخصوصية، بالإضافة إلى متابعة تحركات رؤوس المتحدثين، مع تعديل اتجاه الصوت تبعاً لحركتهم.

وبمشاركة 29 متطوعاً، اختُبر النظام في 10 بيئات مختلفة داخلية وخارجية، مثل الشوارع، والمقاهي، والمتاحف، وأماكن عامة، وأثبت فاعليته في هذه البيئات الواقعية، حتى مع وجود ضوضاء خلفية. وركزت التجربة على 3 لغات هي الإسبانية، والألمانية، والفرنسية، وأظهر النظام أداء جيداً في الترجمة من هذه اللغات إلى الإنجليزية، والعكس.

وأظهرت النتائج أن معظم المستخدمين فضّلوا نظام الترجمة الجديد على الأنظمة التقليدية؛ لما وفره من دقة عالية في الترجمة وتجربة سمعية طبيعية تُحاكي الواقع. كذلك أثبت النظام كفاءته العالية في تتبُّع المتحدثين وتحديد عددهم بدقة، حتى مع تحركهم، مع الحفاظ على اتجاه كل صوت وخصائصه التعبيرية بشكل منفصل.

وقال الباحثون إن النظام يمكّن المستخدمين من فهم ما يقوله عدة أشخاص في الوقت نفسه بلغات مختلفة، في بيئة مزدحمة، مع الحفاظ على أصواتهم الأصلية وتحديد مَن قال ماذا، بما يفتح آفاقاً جديدة للسياحة، والتعليم، والتفاعل الثقافي. ويأمل الفريق بأن يسهم هذا الابتكار، الذي يمكن توسيعه ليشمل نحو 100 لغة، في كسر الحواجز اللغوية بين الثقافات، ليتمكن المستخدم من فهم ما يدور حوله بلغته.

مواضيع الذكاء الاصطناعي | تكنولوجيا | أميركا